



# Building Trust: NDSA Levels of Digital Preservation



U.S. Department of the Interior  
U.S. Geological Survey

**John Faundeen**  
**USGS EROS Center**  
**July 23, 2014**

# Outline

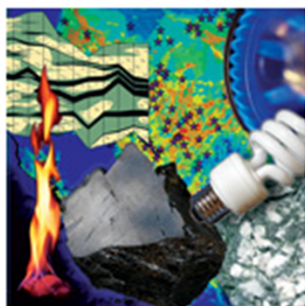
- **Context**
- **NDSA Levels of Digital Preservation**
- **USGS Adoption & Modification**

# Context

- **U.S. Geological Survey**
- **Fundamental Science Practices**
- **Data Preservation Sub-Committee**

## Energy and Minerals Programs

### Energy Resources



The mission of the USGS Energy Resources Program is to understand the processes critical to the formation, accumulation, occurrence, and alteration of geologically based energy resources; to conduct scientifically robust assessments of those resources; and to study the impact of energy resource occurrence and/or production and use on both environmental and human health.

### Mineral Resources



The USGS Mineral Resources Program (MRP) provides scientific information for objective resource assessments and unbiased research results on mineral potential, production, consumption, and environmental effects. The MRP is the sole Federal source for this information.

### USGS Science and Decisions Center



The Science and Decisions Center is a multidisciplinary and open architecture focal point for advancing science-based resource management decisions through the use of USGS scientific information and tools such as ecosystem services, adaptive management, and resilience.

### National Laboratories



The USGS maintains a variety of high quality, state-of-the-art scientific facilities that provide analytical capabilities, scientific support functions, experimental and modeling expertise, and field capabilities. Two of these facilities along with scientific staff are housed in Denver, Colorado at the Federal Center and at Stanford University near the Menlo Park USGS campus. Detailed information on these facilities is available at:

[SHRIMP-RG](#) (Sensitive High-Resolution Ion Microprobe with Reverse Geometry)

[TRIGA](#) (reactor facility)

## **Ecosystems Science**

### **Invasive Species Program**



Provides methodologies and information to reduce threats from invasive species.

### **Cooperative Research Units**



Supports natural resource management decisions through research, education and technical assistance.

### **Fisheries Program**



Supports management, conservation, and restoration of the Nation's fish and aquatic resources.

### **Environments Program**



Conducts ecosystem science and develops models for restoration and adaptive management.

[Priority Ecosystem Science](#)

### **Status & Trends Program**



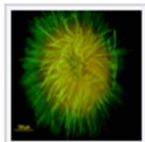
Monitors the Nation's living resources and their habitats.

### **Wildlife Program**



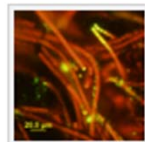
Conducts research on wildlife and their habitats.

### **Genetics & Genomics**



Applies genetics and genomics to answer questions of endangered species, wildlife populations, captive breeding and reintroductions, sources of nonnative species introductions and hybridization and functional genomics.

### **Microbiology**



Research in fish and wildlife health and disease, geomicrobiology, ecosystem function, climate change, water quality for drinking and in recreation, bioremediation, nanotechnology, energy and geographic patterns of microbial distribution.

### **International**



Facilitates synergistic international collaboration because ecoregions and species span and cross boundaries.

### **USA National Phenology Network**



Organizes and centralizes monitoring and scientific information to document changing seasonal patterns.

## Core Science Systems

### Core Science Analytics, Synthesis, and Libraries



With expertise in technology, informatics, and science, CSAS&L leads the management and delivery of scientific data and information for the USGS. CSAS&L implements and promotes standards and best practices to enable efficient, data-driven science for decision making at multiple scales. Critical relationships are established and fostered to identify and access data, and to develop and deploy appropriate technological solutions that support rapid response to emerging natural resource issues. CSAS&L facilitates the [USGS Community for Data Integration](#) (CDI) and supports the [John Wesley Powell Center for Analysis and Synthesis](#), which allows USGS researchers and their colleagues to address complex earth system and natural resource questions in intensive, short-term research focus groups.

### Libraries



The USGS Library program, managed by the Core Science Analytics, Synthesis, and Libraries program, supports all of the fundamental scientific research conducted within the USGS. The library serves both internal and external users with comprehensive access to the literature, data, and information necessary to understand the mission areas of the USGS and make critical decisions about how to proceed with research initiatives and investigations in the earth and natural sciences.

### National Geological and Geophysical Data Preservation Program



The Energy Policy Act of 2005 ([Public Law 109-58](#), Sec. 351) established the National Geological and Geophysical Data Preservation Program in the USGS and outlined the following goals:

- Archive geological, geophysical, and engineering data, maps, well logs, and samples
- Provide a national catalog of archived materials
- Provide technical and financial assistance to State geological surveys and relevant Department

of the Interior bureaus for archived materials

### National Geospatial Program



The National Geospatial Program (NGP) organizes, maintains, and publishes the geospatial baseline of the Nation's topography, natural landscape, and built environment. The baseline is [The National Map](#), a set of databases of map data and information from which customers can download data and derived map products and use web-based map services. Through the [Geospatial Liaison Network](#), the NGP works with cooperators to share the costs of acquiring and maintaining these geospatial data. [The National Atlas of the United States of America®](#), the small-scale component of [The National Map](#), fosters an understanding of broad geographic patterns, trends, and conditions useful for national assessments. The [Federal Geographic Data Committee](#) promotes consistent data and metadata standards, system interoperability, and cross-government best business practices for geospatial resources, policies, standards, and technology as part of the [National Spatial Data Infrastructure](#).

### National Cooperative Geologic Mapping Program



The National Cooperative Geologic Mapping Program (NCGMP) produces accurate geologic maps and 3-D geologic frameworks that provide critical data for sustaining and improving the quality of life and economic vitality of the Nation. [Geologic maps](#) are indispensable to understanding earth surface processes and ground-water availability and quality, supporting DOI land management decisions, mitigating hazards, assisting in ecological and climatic monitoring and modeling, and understanding onshore-offshore sediment processes. NCGMP is unique in the Federal Government as it supports the production of most geologic maps in the United States through a successful [Federal-State-university](#) partnership.

## Climate and Land Use Change Programs

### Land Change Science (LCS)



The goal of the USGS Land Change Science (LCS) Program, formerly Geographic Analysis and Monitoring, is to understand the patterns, processes, and consequences of changes in land use, land condition, and land cover at multiple spatial and temporal scales, resulting from the interactions between human activities and natural systems.

### Land Remote Sensing (LRS)



The Land Remote Sensing Program operates the Landsat satellites and provides the Nation's portal to the largest archive of remotely sensed land data in the world, supplying access to current and historical images. These images serve many purposes from assessing the impact of natural disasters to monitoring global agricultural production.

### National Climate Change and Wildlife Science Center (NCCWSC)



The National Climate Change and Wildlife Science Center responds to the research and management needs of partners and provides science and technical support regarding the impacts of climate change on fish, wildlife and ecological process. The Center is taking the lead on establishing the Department of the Interior regional Climate Science Centers.

### Research and Development Program (R&D)



The USGS Climate and Land Use Change Research and Development Program supports fundamental scientific research to: 1) understand processes controlling Earth system responses to global change over broad temporal and spatial scales; and 2) understand and model impacts of climate and land-cover change on ecosystems and other natural resources.

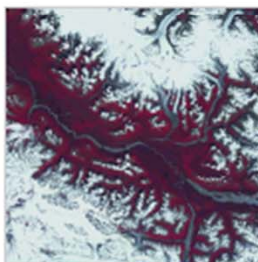
### Carbon Sequestration



sequestration.

Scientists at the U.S. Geological Survey (USGS) are working to assess both the potential capacities and the potential limitations of the various forms of carbon sequestration and to evaluate their geologic, hydrologic, and ecological consequences. In accordance with the Energy Independence and Security Act of 2007, the USGS has developed scientifically based methods for assessment of biologic and geologic carbon

### Earth Resources Observation and Science Center (EROS)



The USGS Earth Resources Observation and Science Center (EROS) contributes to the Climate and Land Use Change Mission Area through research and operational activities that enable the understanding of local to global land change. The EROS multidisciplinary staff uses their unique expertise in remote sensing-based science and technologies to carry out basic and applied research, data acquisition, systems engineering, information access and management, and archive preservation to address the Nation's most critical needs.

## [Coastal and Marine Geology](#)

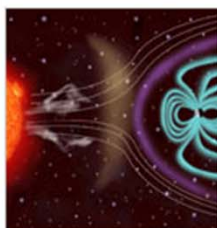


CMGP conducts research on the changes to the coastal and marine environment that impact lands, lives and livelihoods, and vulnerable ecosystems; providing science to inform decisions that ensure safe and resilient coastal communities and sustainable use and protection of marine resources. Supporting the Natural Hazards mission, CMGP conducts research on marine geohazards including earthquakes, tsunami, and submarine landslides and on coastal change hazards from erosion, hurricanes and other extreme storms, and sea-level rise.

- [Coastal Change](#)
- [Tsunamis](#)

[Monthly Newsletter - Soundwaves](#)

## [Geomagnetism](#)



The mission of the Geomagnetism Program is to monitor the Earth's magnetic field. Using ground-based observatories, the Program provides continuous records of magnetic field variations covering long timescales; disseminates magnetic data to various governmental, academic, and private institutions; and conducts research into the nature of geomagnetic variations for purposes of scientific understanding and hazard mitigation.

- [Real-time Data](#)

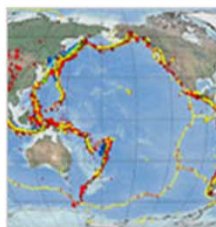
["Monitoring the Earth's Magnetic Field" Fact Sheet](#)

## [Landslide Hazards](#)



The mission of the Landslide Hazards Program is to provide information that leads to the reduction of losses from landslides and increase in public safety through improved understanding of landslide hazards and strategies for hazard mitigation. The USGS conducts landslide hazard assessments, pursues landslide investigations and forecasts, provides technical assistance to respond to landslide emergencies, and engages in outreach activities.

## [Earthquake Hazards](#)



The Earthquake Hazards Program is part of the National Earthquake Hazards Reduction Program (NEHRP), a four-agency partnership led by the National Institute of Standards and Technology (NIST). Earthquakes pose significant risk to 75 million Americans in 39 States. The USGS provides research and information products for earthquake loss reduction, including hazard and risk assessments, comprehensive real-time earthquake monitoring, and public outreach.

- [Earthquake Map](#)
- [Real-time RSS Feeds](#)
- [Earthquake Notification Service](#)
- [Did You Feel It?](#)

["Earthquake Hazards – A National Threat" Fact Sheet](#)

## [Global Seismographic Network](#)



The Global Seismographic Network (GSN) is a permanent digital network of state-of-the-art seismological and geophysical sensors connected by a telecommunications network. The GSN provides near-uniform, worldwide monitoring of the Earth, with over 150 modern seismic stations distributed globally. The GSN was formed in partnership among the USGS, the National Science Foundation (NSF) and the Incorporated Research Institutions for Seismology (IRIS).

- [Station Status](#)

[The Global Seismographic Network Fact Sheet](#)

## [Volcano Hazards](#)



The Volcano Hazards Program advances the scientific understanding of volcanic processes in order to lessen the harmful impacts of volcanic activity. The USGS monitors active and potentially active volcanoes, assesses their hazards, responds to volcanic crises, and conducts research on how volcanoes work. The USGS also issues warnings of potential volcanic hazards to responsible emergency-management authorities and to the populace affected.






## Streams, Lakes, and Reservoirs

Science to observe, analyze, and understand the movement and condition of surface water.

- ◆ [USGS Surface Water Information](#)
- ◆ [USGS Flood Information](#)
- ◆ [Flood inundation mapping](#)
- ◆ [Estimating streamflow statistics](#)
- ◆ [Suspended sediment](#)
- ◆ [Watershed modeling](#)



## Groundwater, Aquifers, and Wells

Science to discover and describe the location, condition, and behavior of water in the ground.

- ◆ [USGS Groundwater information](#)
- ◆ [What is groundwater?](#)
- ◆ [Groundwater and drought](#)
- ◆ [How do I find USGS groundwater information?](#)
- ◆ [Selected USGS groundwater publications](#)



## Quality of Water Resources

Science to monitor and evaluate biological, chemical, and environmental factors affecting water quality.

- ◆ [USGS Water Quality information](#)
- ◆ [What is water quality?](#)
- ◆ [Popular water-quality topics and issues](#)
- ◆ [Selected water-quality publications](#)



## Water Use

Information about how water is used now and in the past.

- ◆ [National Water Use home page](#)
- ◆ [2005 National water-use publication](#)
- ◆ [Download water-use data](#)
- ◆ [Publications](#)
- ◆ [Water-Science School: Water use](#)

## Environmental Health Programs

### Contaminant Biology



The USGS Contaminant Biology Program investigates the effects and exposure of environmental contaminants to the Nation's living resources, particularly those under the stewardship of the Department of the Interior.

### Toxic Substances Hydrology



The USGS Toxic Substances Hydrology Program provides objective scientific information on environmental contamination to improve characterization and management of contaminated sites, to protect human and environmental health, and to reduce potential future

contamination problems.

# NDSA Levels of Digital Preservation

**“The goal of this document is to provide a basic tool for helping organizations manage and mitigate digital preservation risks.”**

**“This is useful for developing plans — not a plan in itself”**

**“These levels are non-judgmental: Organizations have different resources and priorities, and as a result need to think about how to best allocate those resources to meet their specific needs.”**

**“These levels can be applied to collection(s) or system(s): These levels function coherently with everything from individual case by case collection level decisions as well as issues for an entire centralized repository”**



# NDSA Levels of Digital Preservation

	<b>Level One (Protect Your Data)</b>	<b>Level Two (Know Your data)</b>	<b>Level Three (Monitor Your Data)</b>	<b>Level Four (Repair Your Data)</b>
<b>Storage and Geographic Location</b>	<ul style="list-style-type: none"> <li>• Two complete copies that are not collocated</li> <li>• For data on heterogeneous media (optical disks, hard drives, etc.) get the content off the medium and into your storage system</li> </ul>	<ul style="list-style-type: none"> <li>• At least three complete copies</li> <li>• At least one copy in a different geographic location</li> <li>• Document your storage system(s) and storage media and what you need to use them</li> </ul>	<ul style="list-style-type: none"> <li>• At least one copy in a geographic location with a different disaster threat</li> <li>• Obsolescence monitoring process for your storage system(s) and media</li> </ul>	<ul style="list-style-type: none"> <li>• At least 3 copies in geographic locations with different disaster threats.</li> <li>• Have a comprehensive plan in place that will keep files and metadata on currently accessible media or systems</li> </ul>
<b>File Fixity and Data Integrity</b>	<ul style="list-style-type: none"> <li>• Check file fixity on ingest if it has been provided with the content</li> <li>• Create fixity info if it wasn't provided with the content</li> </ul>	<ul style="list-style-type: none"> <li>• Check fixity on all ingests Use write-blockers when working with original media</li> <li>• Virus-check high risk content</li> </ul>	<ul style="list-style-type: none"> <li>• Check fixity of content at fixed intervals</li> <li>• Maintain logs of fixity info; supply audit on demand</li> <li>• Ability to detect corrupt data</li> <li>• Virus-check all content</li> </ul>	<ul style="list-style-type: none"> <li>• Check fixity of all content in response to specific events or activities</li> <li>• Ability to replace/repair corrupted data</li> <li>• Ensure no one person has write access to all copies</li> </ul>



# NDSA Levels of Digital Preservation

<b>Information Security</b>	<ul style="list-style-type: none"> <li>• Identify who has read, write, move, and delete authorization to individual files</li> <li>• Restrict who has those authorizations to individual files</li> </ul>	<ul style="list-style-type: none"> <li>• Document access restrictions for content</li> </ul>	<ul style="list-style-type: none"> <li>• Maintain logs of who performed what actions on files, including deletions and preservation actions</li> </ul>	<ul style="list-style-type: none"> <li>• Perform audit of logs</li> </ul>
<b>Metadata</b>	<ul style="list-style-type: none"> <li>• Inventory of content and its storage location</li> <li>• Ensure backup and non-collocation of inventory</li> </ul>	<ul style="list-style-type: none"> <li>• Store administrative metadata</li> <li>• Store transformative metadata and log events</li> </ul>	<ul style="list-style-type: none"> <li>• Store standard technical and descriptive metadata</li> </ul>	<ul style="list-style-type: none"> <li>• Store standard preservation metadata</li> </ul>
<b>File Formats</b>	<ul style="list-style-type: none"> <li>• When you can give input into the creation of digital files encourage use of a limited set of known open file formats and codecs</li> </ul>	<ul style="list-style-type: none"> <li>• Inventory of file formats in use</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor file format obsolescence issues</li> </ul>	<ul style="list-style-type: none"> <li>• Perform format migrations, emulation and similar activities as needed</li> </ul>



# USGS Adoption & Modification

- *USGS Guidelines for the Preservation of Digital Scientific Data*